



A Hybrid Double-Observer Sightability Model for Aerial Surveys

PAUL C. GRIFFIN,¹ U.S. Geological Survey, Forest and Rangeland Ecosystem Science Center, 600 E. Park Ave., Port Angeles, WA 98362, USA

BRUCE C. LUBOW, Colorado State University, Fort Collins, CO, USA

KURT J. JENKINS, U.S. Geological Survey, Forest and Rangeland Ecosystem Science Center, 600 E. Park Ave., Port Angeles, WA 98362, USA

DAVID J. VALES, Muckleshoot Indian Tribe, Auburn, WA, USA

BARBARA J. MOELLER, Puyallup Tribe of Indians, Puyallup, WA, USA

MASON REID, Mount Rainier National Park, Ashford, WA, USA

PATRICIA J. HAPPE, Olympic National Park, Port Angeles, WA, USA

SCOTT M. MCCORQUODALE, Washington Department of Fish and Wildlife, Yakima, WA, USA

MICHELLE J. TIRHI, Washington Department of Fish and Wildlife, Lakewood, WA, USA

JIM P. SCHABERL, Shenandoah National Park, Luray, VA, USA

KATHERINE BEIRNE, North Coast and Cascades Network Inventory and Monitoring Program, Olympic National Park, Port Angeles, WA, USA

ABSTRACT Raw counts from aerial surveys make no correction for undetected animals and provide no estimate of precision with which to judge the utility of the counts. Sightability modeling and double-observer (DO) modeling are 2 commonly used approaches to account for detection bias and to estimate precision in aerial surveys. We developed a hybrid DO sightability model (model M_H) that uses the strength of each approach to overcome the weakness in the other, for aerial surveys of elk (*Cervus elaphus*). The hybrid approach uses detection patterns of 2 independent observer pairs in a helicopter and telemetry-based detections of collared elk groups. Candidate M_H models reflected hypotheses about effects of recorded covariates and unmodeled heterogeneity on the separate front-seat observer pair and back-seat observer pair detection probabilities. Group size and concealing vegetation cover strongly influenced detection probabilities. The pilot's previous experience participating in aerial surveys influenced detection by the front pair of observers if the elk group was on the pilot's side of the helicopter flight path. In 9 surveys in Mount Rainier National Park, the raw number of elk counted was approximately 80–93% of the abundance estimated by model M_H . Uncorrected ratios of bulls per 100 cows generally were low compared to estimates adjusted for detection bias, but ratios of calves per 100 cows were comparable whether based on raw survey counts or adjusted estimates. The hybrid method was an improvement over commonly used alternatives, with improved precision compared to sightability modeling and reduced bias compared to DO modeling. © 2013 The Wildlife Society.

KEY WORDS aerial survey, *Cervus elaphus*, double-observer, elk, heterogeneity bias, Mount Rainier, sightability model.

Uncorrected counts of animals observed during aerial wildlife surveys generally underestimate true abundance and misrepresent the composition of populations because of detection bias—the failure to detect all individuals or groups of animals present during a survey (Samuel and Pollock 1981, McCorquodale 2001). Methods for estimating detection bias in aerial surveys include sightability models (Steinhorst and Samuel 1989), distance models (Burnham and Anderson 1984, Buckland et al. 2004), mark-resight models (White 1996, Skalski et al. 2005) including double-observer (DO) models (Graham and Bell 1989), and methods that combine 2 or more such techniques (Quang and Becker 1997, Buckland et al. 2010).

Sightability models (model M_S) are commonly based on logistic regression equations used to predict the probability that aerial observers detect individuals or groups of animals as functions of sighting covariates such as group size, the amount of concealing vegetation near the group, and animal movement at the time of detection (Samuel et al. 1987, Unsworth et al. 1999, Gilbert and Moeller 2008, McIntosh et al. 2009). Model M_S has been used extensively to estimate observation biases of aerial surveys for elk (*Cervus elaphus*; Samuel et al. 1987, Anderson et al. 1998, McCorquodale 2001, Gilbert and Moeller 2008, McIntosh et al. 2009), moose (*Alces alces*; Anderson and Lindzey 1996, Giudice et al. 2012), mountain sheep (*Ovis dalli*; Udevitz et al. 2006) mountain goats (*Oreamnos americanus*; Rice et al. 2008), and other species (Krueger et al. 2007, Manning and Garton 2011). Model M_S is fit to records of detection or non-detection of known animal groups present during surveys

Received: 24 January 2013; Accepted: 19 June 2013

Published: 19 September 2013

¹E-mail: pgriffin@usgs.gov

that are identified using radio-telemetry or simultaneous observation by independent crews. Sighting covariates are recorded for groups that are detected by the aerial survey crew during surveys and for known missed groups. Based on the assumptions that the known monitored animals are distributed at random through the population being sampled and that covariates are recorded accurately, sightability trials reflect the unconditional, unbiased effect of covariates on detection probability (Steinhorst and Samuel 1989) at the time of the sightability trials. Estimates of detection probabilities permit the computation of group-specific correction factors used to account for detection bias. Once an M_S model has been developed, it is typically applied to future surveys without further reliance on radio-marked animals, which raises an important caveat: one must assume that the conditions under which the model was developed continue to characterize future surveys (Williams et al. 2002). Another disadvantage is that, because known marked groups are typically a small subset of all available groups, M_S models are usually calibrated by fitting to a relatively small number of data points. In contrast, a method that quantifies detection probability using all observations from ongoing surveys would provide larger sample sizes and be more robust to temporal changes that may affect how survey conditions influence detection probability.

Simultaneous DO sampling also has been used extensively to estimate detection biases for a wide variety of species when marked or radio-collared animals are not available (Caughley and Grice 1982, Graham and Bell 1989, Crête et al. 1991, Pollock et al. 2006, Shirley et al. 2012). DO models (model M_D) use observations recorded by ≥ 2 independently searching observers to estimate the overall probability that any given group of animals is detected by ≥ 1 of the observers. Data from every observed group on every survey contribute to estimating sighting probabilities and correcting for unseen animals. This potentially large sample of observed animal groups available for model M_D helps to improve precision and reduce variance of population estimates (Marsh and Sinclair 1989, Potvin et al. 2004).

Early DO methods estimated abundance using simple mark-resighting estimators (e.g., Lincoln-Petersen), which may be biased by heterogeneity in visibility among groups of animals (Pollock and Kendall 1987). The effect of heterogeneity (Southwell et al. 2007) can be reduced by developing models of sighting probability that account for effects of recorded sighting covariates on detection probabilities. DO data with individual covariates for each animal group can be analyzed using a conditional likelihood approach (Huggins 1989, 1991) with logistic regression models of sighting probability (Walter and Hone 2003). However, sightings of animal groups by each observer are only conditionally independent because DO data sets are limited to animals that were seen by ≥ 1 of the observers. Even accounting for sighting covariates, unknown and unmodeled sources of heterogeneity may cause some groups with similar covariate values to be more likely to be seen by both observers or to have been missed by both observers because of some unknown and unmeasured effect (Southwell

et al. 2007, Barker 2008, Laake et al. 2011). As a result, a disadvantage of using only DO analyses is that detection probabilities may be biased high (Barker 2008), resulting in abundance underestimation (Caughley and Grice 1982). Bias can be reduced substantially using well-chosen covariates, but it is impossible to know, measure, and correct for all possible sources of sighting differences explicitly, thereby leaving at least some residual unmodeled heterogeneity and resulting bias (heterogeneity bias). However, bias caused by unmodeled heterogeneity can be quantified if one has another source of independent observations of randomly selected groups of animals.

To address the shortcomings of the M_S and M_D models, we developed a candidate set of hybrid DO sightability models (M_H models) for model fitting and multi-model inference (Burnham and Anderson 2002). These M_H models are structured to include effects of specified covariates on the unconditional probability that a given pair of observers in the helicopter (front-seat pair or back-seat pair) detects a given elk group. The M_H models include parameters that account for sighting covariates and for the residual heterogeneity bias that is not accounted for by the explicitly measured covariates. The M_H models are fit to data from 2 sources: 1) double-observer sightability (DO-S) trials with 2 independent observers and ≥ 1 radio-marked animals in the group; and 2) DO observations of elk groups with no radio-marked animals in the group. The DO-S trials provide independent data used to estimate the implicit residual heterogeneity bias that would not be quantifiable from DO observations alone.

We demonstrate an application of M_H models with elk surveys in high-elevation summer ranges in Mount Rainier National Park, Washington, USA. We parameterized candidate suites of M_H models, standard M_S models, and M_D models with sighting covariates. We compared estimates of group-specific detection probabilities from M_H models and M_D models to assess the level of bias resulting from unmodeled heterogeneity in DO methods. We also compared survey results obtained from M_H , M_S , and M_D modeling approaches. We hypothesized that the M_D approach would overestimate detection probabilities and, consequently, underestimate abundance relative to the M_H and M_S approaches. We also predicted that estimates from model M_H would be more precise than those from model M_S because of the larger data set available to parameterize the M_H models.

STUDY AREA

The study area within Mount Rainier National Park (NP) was defined by summer ranges of 2 elk herds that migrate annually to the northeast and southeast slopes of Mount Rainier in Washington's Cascades Range (Fig. 1). The North Rainier herd survey area (103 km²) was bounded by elevations below 2,100 m and above 1,500 m, whereas the South Rainier herd survey area (89 km²) was bounded by elevations below 2,100 m and above 1,350 m except on some southwest-facing slopes where past landslides and wildfires maintained open parklands down to 1,200 m. We referred to

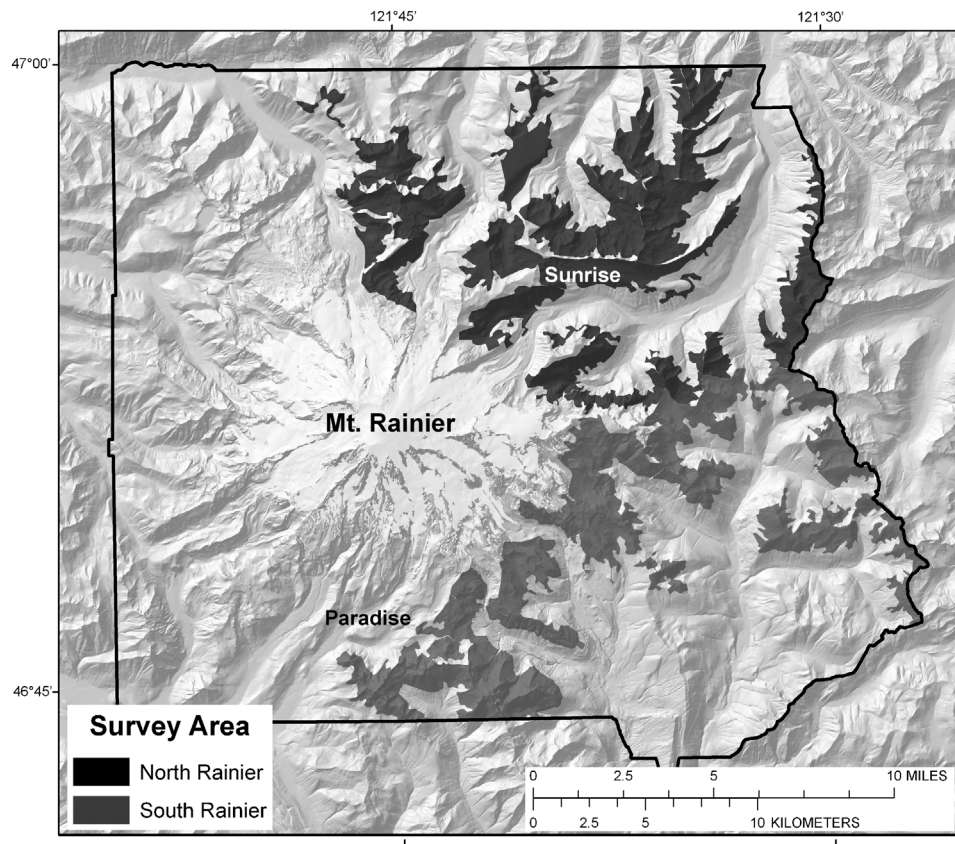


Figure 1. Elk survey areas within Mount Rainier National Park, Washington, USA, 2008–2010.

satellite-derived estimates of vegetation cover and type (Pacific Meridian Resources 1996) and excluded areas of continuously dense forest canopy cover, rock, or permanent snow. Because the study area did not encompass a geographically closed elk population, our scope of inference is limited to estimating abundance of elk using subalpine parklands during summer. Trends in the number of elk using these areas, and their sex and age composition, are of interest to National Park Service (NPS), state, and tribal wildlife managers because of past NPS concerns about overabundance of elk on summer ranges in the park and the utility of these indices for managing harvests of elk outside the park to meet multiple agency or tribal wildlife management objectives. The study area encompassed subalpine parklands of the upper mountain hemlock (*Tsuga mertensiana*) forest zone consisting of mosaics of forest patches, tree clumps, and subalpine meadows (Franklin et al. 1988). The tallest trees in the subalpine parklands were <20 m. Canopy closure was highly variable, but tended not to be dense near treeline.

METHODS

Aerial Surveys

We surveyed elk in each survey area by helicopter (Bell 206BIII, Bell Helicopter, Hurst, TX or Hughes 500D, MD Helicopters, Inc., Mesa, AZ) between 15 August and 15 September 2008–2010. Surveys began as early as 4 hours

before local sunset and ended as late as 30 minutes after sunset. Helicopter flights were approximately 100–150 m above ground level with flight lines approximately 250–500 m apart. Flight speed was approximately 85 km/hr with a search intensity of roughly 2.8 min/km². Typically, we surveyed a single survey area using 2 helicopters and survey crews. We divided each survey area into subunits to improve survey logistics, communications, and safety. We conducted 1–2 complete surveys of each survey area in each year for a total of 9 surveys. To increase the number of DO-S trials used for model development, we conducted additional survey flights in subjectively chosen subunits that were likely to contain radio-marked elk. For these additional flights, the survey crews did not know which subunits would contain collared elk.

A pilot and 3 experienced observers conducted surveys. The pilot and front-seat observer had views to the front, sides, and below the flight path. Each back-seat observer had a view to 1 side of the flight path. Although pilots' attention was primarily on flying safely, pilots also detected elk groups. In-flight procedures required each observer to search independently for elk groups in the survey area, but to collaborate in determining group size, composition, and covariates of detected groups after each flight member had an opportunity to detect a given elk group (e.g., Ransom 2012). We noted the rare occasions when either observer pair was inadvertently alerted about the presence of an elk group by the other before having a chance to record an observation

independently, so we omitted those non-independent observations from the data set used for model fitting. We recorded the following covariates associated with each group: group size and composition, percent concealing vegetation ≤ 10 m from the group (V ; within 5 classes: 0, 1–25, 26–50, 51–75, or 76–100%), whether the group was standing on herbaceous vegetation (H ; yes or no), whether the group was in forest vegetation (F ; yes or no), lighting conditions (L ; flat or high contrast), and the activity of the group when first detected (M ; moving or not). We noted whether each elk group was detected to the left, right, or both sides of the helicopter's flight path. We noted when an elk group was directly below the helicopter's flight path and not visible to the back-seat observers (C ; centerline yes or no). A separate covariate denoted whether the elk group was on the same side of the helicopter's flightline as an inexperienced pilot-observer (P ; yes if group on same side as a pilot-observer with little experience conducting aerial wildlife surveys, no if pilot was experienced in conducting aerial wildlife surveys or elk group was on other side). We photographed groups with ≥ 20 individuals to reduce bias in group size estimation (Cogan and Diefenbach 1998, Schoenecker et al. 2006). After the flight, we inspected multiple photographs from each large group and updated the group size and composition as needed.

A sample of female elk from both the North and South Rainier herds was radio-collared prior to the surveys as part of companion studies (D. Vales, Muckleshoot Indian Tribe, unpublished data; Moeller 2010). In 2008, 2009, and 2010, the surveyed areas contained 20, 29, and 22 collared elk, respectively. Because we had not collared male elk for use in model development, application of the resulting models to both male and female elk is based on an assumption that there was no interaction between sex and the effects of sighting covariates on detection probability. Although we assume identical parameters for estimating sighting probability of male and female elk, groups that are primarily male and groups that are primarily female tend to have different observed values for sighting covariates such as group size, and percent concealing vegetation (McCorquodale 2001). In other words, the models are the same for all groups but the inputs typically differ, so the models predict different sighting probabilities. For every elk group detected in visual surveys, 1 observer in the helicopter used a radio-telemetry receiver to determine whether any radio-collared elk were present in the group and recorded their identity. After completing a survey within a subunit or small group of subunits, we used radio-telemetry to identify and locate any radio-collared elk that were missed during the surveys and we recorded covariates for those missed groups.

If either the pilot or the front-seat observer detected an elk group, we recorded the group as seen by the front observer pair. If either of the back-seat observers detected an elk group, we recorded it as seen by the back observer pair. If a radio-collared animal was located in the elk group using radio-telemetry, we recorded it as detected by radio (hereafter referred to as the radio observer).

Model Development

Variable treatments.—Before model fitting, we transformed covariates. We converted group size to the natural log of group size, $\ln(n)$, and V to the decimal midpoint of the recorded range. We converted H , F , L , M , C , and P to indicator variables with values of 1 or 0. We imputed missing covariate values for groups containing collared elk that we missed during surveys and for which we could not determine covariate values in the field. This was necessary because we could not always determine group sizes or group activity beneath dense tree cover. For 9 DO-S trials with missing data for group size and activity, we substituted the median group size computed from all groups that contained ≥ 1 cow and were detected in the same category of percent concealing vegetation. For 1 DO-S trial with no percent concealing vegetation recorded, we used the median group size and vegetation cover score based on all observed groups with ≥ 1 cow. For the missing activity records, we used the modal activity based on all groups with ≥ 1 cow. We used audio recordings of the cockpit conversation and global positioning system (GPS) records of the survey flight path to determine on which side of the helicopter an elk group was located when that information was not recorded on paper in flight. For DO-S trials, we used GPS and a geographic information system (GIS) to determine on which side of the helicopter's flight path elk groups were located when the elk group's location relative to the helicopter's flight line was not recorded. For 3 DO-S trials for which we could not determine side-of-helicopter from the recorded by crew members on both the right and left sides of the helicopter. Because the number of DO observations was much larger than the number of DO-S trials, we did not impute covariate values, and omitted from model creation any DO observations for which we had a missing value for any sighting covariate or the record of the side of the helicopter's flight path on which the elk group was located.

We defined sighting probabilities, $\hat{p}_{g,b,j,m(k)}$, where g indicates the group type (U =unconditional, C =conditional) explained below, b is the observer (F =front, B =back, or R =radio), j indexes the elk group (1, 2, ..., n , where n is the number of groups detected), and $m(k)$ is the model structure where m indicates the model type (H for M_H models, S for M_S models, and D for M_D models), and k indexes the hypothesized alternative model parameterization of each type ($k=1, 2, \dots, M$ where M is the number of model parameterizations for that model type).

Hybrid double-observer sightability models.—We compiled detection histories for all observed elk groups, denoting whether front-seat, back-seat, or radio-telemetry observers detected each elk group in the data set. Each elk group belonged to 1 of 2 categories:

1. Groups available for detection by all 3 observers (front, back, and radio); these DO-S trials contained ≥ 1 radio-collared elk. Animal groups in this category were always detected by telemetry so long as they were in the surveyed area. A group's unconditional probability of detection should not depend on whether it includes a radio-collared

animal. The estimated detection probabilities $\hat{p}_{U,b,j,H(k)}$ are unconditional (indexed by $g = U$) because marked elk groups are all detected by telemetry, and their inclusion in the data set is not contingent on any circumstance other than that the group was present in the surveyed area; these are presumed to be a random sample of all elk groups.

2. Groups available for detection by only the front and back observers; these DO observations contained no radio-collared elk. Animal groups in this category were in the data set only if they were in the surveyed area and were detected by either the front- or back-seat observers. The estimated conditional detection probabilities $\hat{p}_{C,b,j,H(k)}$ are likely to be biased because the unmarked elk are in the data set, conditional on being seen by at least 1 observer pair. We refer to these groups as conditional and index them with $g = C$ in our notation.

The inherent difference between $\hat{p}_{U,b,j,H(k)}$ and $\hat{p}_{C,b,j,H(k)}$ is the result of unmodeled heterogeneity that is not adequately explained by other model covariates. This is the degree of heterogeneity bias that would result from using DO trials only. We quantified this degree of bias by including an additive heterogeneity bias parameter in the model structures in the functions for estimating $\hat{p}_{U,F,j,H(k)}$ and $\hat{p}_{U,B,j,H(k)}$. These are the estimated probabilities for an aerial observer seeing a radio collared group, which is, on average, less likely than seeing a group seen by the other aerial observer; thus, the heterogeneity parameter is a negative coefficient that accounts for the difference in detection probabilities related to conditional (nonrandom) DO sampling only versus the unconditional (random) sampling of radio-marked groups.

Because the data represent 2 categories of elk groups and 3 observers, any candidate M_H model has 6 detection probabilities (Table 1). The radio-telemetry observer has perfect detection of unconditional groups ($p_{U,R,j,H(k)} = 1$) and no chance of detecting conditional groups ($p_{C,R,j,H(k)} = 0$). The 4 remaining sighting probabilities are estimated for category g , observer b , elk group j , according to model structure k by the logistic function (Huggins 1989, 1991):

$$\hat{p}_{g,b,j,H(k)} = \frac{e^{\sum_i \hat{\beta}_{k,i} \cdot D_{g,b,H(k),i} \cdot x_{j,i}}}{1 + e^{\sum_i \hat{\beta}_{k,i} \cdot D_{g,b,H(k),i} \cdot x_{j,i}}} \quad (1)$$

for $\{g : U, C\}; \{b : F, B\}; \{j : 1, 2, \dots, n\}; \{k : 1, 2, \dots, m\}$

where $\hat{\beta}_{k,i}$ is a fitted coefficient for model k , where $i = 1, 2, \dots, p$ indexes the list of all p candidate coefficients in all models.

$D_{g,b,k,i}$ is an indicator for model k specific to elk groups in category g , observer b . The value is set to 1 if parameter i is included in model k , and 0 otherwise. $x_{j,i}$ is the observed value of sighting covariate i for elk group j .

We parameterized the models so that β coefficients (Table 2) corresponding to the effects of sighting covariates on front- and back-seat detection probabilities applied to both elk group categories (i.e., with and without collared elk; Table 1). Thus, we estimated sightability coefficients for front- and back-seat observer pairs based on sightability data from all elk groups irrespective of whether or not collared elk were present. The heterogeneity bias parameters, though, could only be estimated with reference to detection patterns of elk groups with ≥ 1 radio-collar. We structured models that contained separate estimates of intercept parameters for front- and back-seat observers as well as models that contained a common intercept for front- and back-seat observer pairs. Similarly, we structured some models with separate heterogeneity bias parameters for front- and back-seat observers, and other models with a single heterogeneity parameter for both pairs of observers combined (Table 3).

After having estimated all parameters for each model, we accounted for heterogeneity bias and computed the unconditional detection probabilities, $\hat{p}_{U,b,j,H(k)}$, expected under each model, $H(k)$, for all observed groups, including groups with no radio-collars. For each candidate model, k , we computed the expected unconditional overall probability that ≥ 1 observer would detect any j given group, $\hat{p}_{U,,j,H(k)}$ (dot notation for the position index indicates overall sighting probability for all observers) as 1 minus the expected probability that it was missed by all observers:

$$\hat{p}_{U,,j,H(k)} = 1 - (1 - p_{U,R,j,H(k)}) \cdot (1 - \hat{p}_{U,F,j,H(k)}) \cdot (1 - \hat{p}_{U,B,j,H(k)}) \quad (2)$$

Groups with ≥ 1 radio collared elk had a fixed probability of $p_{U,R,j,H(k)} = 1$, so their $\hat{p}_{U,,j,H(k)} = 1$.

We structured 22 candidate M_H models reflecting a priori expectations of covariate effects on detection probability as well as different model structures. All candidate M_H models were structured to solve for the 6 detection probabilities (Table 1), but they varied according to which sighting covariates were included, whether intercepts were estimated for front- and back-observers separately or together, and whether heterogeneity bias parameters were estimated for front and back observers separately or together (Tables 2 and 3). We did not consider any M_H model without the heterogeneity bias parameter because we found little support for any such models (difference in corrected Akaike's Information Criterion [ΔAIC_c] > 22). In addition, we decided a priori, based on prior experience, to include group

Table 1. Modeled probabilities of detecting elk group j , within category g (rows), and for observer b (columns), in helicopter surveys of elk in Mount Rainier National Park, Washington, USA, 2008–2010.

Elk group category, g	Observer, b		
	R (Radio-telemetry)	F (Front pair)	B (Back pair)
$g = U$. Radio-collar(s) in group (unconditional)	$\hat{p}_{U,R,j,H(k)} = 1.0$	$\hat{p}_{U,F,j,H(k)}$	$\hat{p}_{U,B,j,H(k)}$
$g = C$. No radio-collar in group (conditional)	$\hat{p}_{C,R,j,H(k)} = 0$	$\hat{p}_{C,F,j,H(k)}$	$\hat{p}_{C,B,j,H(k)}$

Table 2. Model components potentially included in model structures for detection probabilities of elk groups in aerial surveys in Mount Rainier National Park, Washington, USA, 2008–2010.

Model parameter	Description of model parameter
$\beta_{F=B}$	Common intercept for front- and back-observer detection probabilities $p_{g,F,j,m(k)}$ and $p_{g,B,j,m(k)}$.
β_F, β_B	Independent intercepts for front-pair (F) and back-pair (B) observer detection probabilities.
β_D	Common parameter for heterogeneity detection bias effect on detection probabilities $p_{g,F,j,H(k)}$ and $p_{g,B,j,H(k)}$.
β_{D-F}, β_{D-B}	Independent parameters for separate heterogeneity detection bias effects for front-pair (F) and back-pair (B) observers.
$\beta_{\ln(n)}$	Parameter for the effect of the natural logarithm of elk group size on detection probabilities $p_{g,F,j,m(k)}$ and $p_{g,B,j,m(k)}$.
β_P	Parameter for the effect of inexperienced pilots on detection probabilities for front observer, $p_{g,F,j,m(k)}$, when the elk group was on the same side of the helicopter as the pilot.
β_C	Parameter for the effect of the elk group being directly under the helicopter flight path's centerline on back observer probabilities, $p_{g,B,j,m(k)}$.
β_V	Parameter for the effect of the percent concealing vegetation on detection probabilities $p_{g,F,j,m(k)}$ and $p_{g,B,j,m(k)}$.
β_L	Parameter for the effect of flat lighting level on detection probabilities $p_{g,F,j,m(k)}$ and $p_{g,B,j,m(k)}$.
β_M	Parameter for the effect of moving animal activity on detection probabilities $p_{g,F,j,m(k)}$ and $p_{g,B,j,m(k)}$.
β_H	Parameter for the effect of predominantly herbaceous vegetation on detection probabilities $p_{g,F,j,m(k)}$ and $p_{g,B,j,m(k)}$.
β_F	Parameter for the effect of predominantly forest vegetation on detection probabilities $p_{g,F,j,m(k)}$ and $p_{g,B,j,m(k)}$.

Table 3. Candidate models for detection probabilities in elk aerial surveys in Mount Rainier National Park, Washington, USA, 2008–2010. We show the model number and description, number of parameters (K), difference in corrected Akaike's Information Criterion (ΔAIC_c), and Akaike weight ($\hat{w}_{m(k)}$), for model type m , parameterization variant (k) for the 22 hybrid double-observer sightability (M_H) candidate models fit to 570 observations, 6 candidate sightability models (M_S) fit to 97 sightability trials, and 5 candidate double-observer models (M_D) fit to 570 observations. We ranked models from high to low, by ΔAIC_c . We made multi-model inferences using models with $\Delta AIC_c < 4.0$. Parameters included in each model are indicated by an X in the corresponding column. Parameters that are not applicable to the given model type are indicated by —. Parameters that were not considered for a particular model type because they received minimal support in the M_H analysis are indicated with an *.

Model name	Parameter												K	ΔAIC_c	$\hat{w}_{m(k)}$	
	$\beta_{F=B}$	β_F, β_B	β_D	β_{D-F}, β_{D-B}	$\beta_{\ln(n)}$	β_P	β_C	β_V	β_L	β_M	β_F	β_H				
M_H models																
H(1)	X		X		X	X	X	X	X					7	0.00	0.149
H(2)	X		X		X	X	X	X						6	0.72	0.104
H(3)	X		X		X	X	X	X		X				7	0.82	0.099
H(4)		X		X	X	X	X	X	X					9	0.99	0.091
H(5)		X	X		X	X	X	X	X					8	1.19	0.082
H(6)		X		X	X	X	X	X						8	1.70	0.064
H(7)	X		X		X	X	X	X				X		7	1.72	0.063
H(8)		X		X	X	X	X	X		X				9	1.77	0.061
H(9)		X	X		X	X	X	X						7	1.91	0.057
H(10)		X	X		X	X	X	X		X				8	2.02	0.054
H(11)		X		X	X	X	X	X				X		9	2.70	0.039
H(12)	X		X		X	X	X	X			X			7	2.72	0.038
H(13)		X	X		X	X	X	X				X		8	2.92	0.035
H(14)		X		X	X	X	X	X			X			9	3.71	0.023
H(15)		X	X		X	X	X	X			X			8	3.91	0.021
H(16)	X		X		X	X	X					X		6	4.81	0.013
H(17)	X		X		X	X	X				X			6	8.21	0.002
H(18)	X		X		X	X	X		X					6	9.66	0.001
H(19)	X		X		X	X	X							5	9.91	0.001
H(20)	X		X		X	X	X			X				6	10.33	0.001
H(21)		X		X	X	X	X							7	10.65	0.001
H(22)		X	X		X	X	X							6	10.87	0.001
M_S models																
S(1)	X	—	—	—	X		—	X	X					4	0.00	0.30
S(2)	X	—	—	—	X		—	X						3	0.25	0.26
S(3)	X	—	—	—	X		—	X				X		4	1.47	0.14
S(4)	X	—	—	—	X		—	X			X			4	2.13	0.10
S(5)	X	—	—	—	X	X	—	X						4	2.18	0.10
S(6)	X	—	—	—	X		—	X		X				4	2.34	0.09
M_D models																
D(1)	X	*	—	—	X	X	X	X						5	0.00	0.28
D(2)	X	*	—	—	X	X	X	X		X				6	0.21	0.25
D(3)	X	*	—	—	X	X	X	X	X					6	0.44	0.22
D(4)	X	*	—	—	X	X	X	X				X		6	1.17	0.15
D(5)	X	*	—	—	X	X	X	X			X			6	2.01	0.10

size, $\ln(n)$, pilot-observer experience influence on detection probabilities of front-seat observers, P , and position of elk groups directly under the center of the helicopter flight path influence on back-seat observer detection probabilities, C , in all M_H models. To avoid overfitting, we limited model structures to include ≤ 2 additional sighting covariates from among the remaining candidates.

Sightability models.—For comparison with the M_H model, we developed sightability models (M_S) fit to data from groups containing radio-collared elk (Steinhorst and Samuel 1989). These use the same logistic structure as M_H models (Equation 1). Because M_S models are based on the collective observations of the aerial survey crew, we did not include information about the DO pattern (i.e., front and rear) of independent detections in M_S model development. The M_S models had no heterogeneity bias parameter because all data are from the unconditional (radio collared) groups. Thus, of the 6 sighting probabilities (Table 1), only 2 are applicable here: $\hat{p}_{U,b,j,S(k)}$, where b was either R for radio observer or A for the pooled aerial observers, and only 1 of these is estimated because $\hat{p}_{U,R,j,S(k)} = 1.0$. We fit 6 Huggins (1989, 1991) models in Program MARK (White and Burnham 1999), fixing the probability for telemetry observer equal to 1. The candidate M_S models all included an intercept, effects of $\ln(n)$ and V , and at most 1 other covariate that was highly supported in the M_H models (F , H , L , M , and P ; see Results) because of limited sample size.

Double-observer models.—We parameterized 5 Huggins (1989, 1991) DO models (M_D) with individual covariates fit to our DO observation data in Program MARK (White and Burnham 1999). The basic structure of these also follows Equation (1). Because our intent was to compare M_D model performance to that of the M_H model, we did not include any detections based on radio-telemetry. M_D models included an intercept, $\ln(n)$, C , V , and P plus up to 1 more of the most highly supported covariates from the M_H models considered (F , H , L , and M ; see Results). M_D models included no heterogeneity parameter because all observations were conditional. Thus, only 2 sighting probabilities were applicable: $\hat{p}_{C,b,j,D(k)}$, where b was either F for front or B for the back aerial observers and both of these were estimated from the data.

Probability of Detection

To assess heterogeneity bias effects in the M_H model across a range of sighting covariate, $x_{j,i}$ values, we computed estimated unconditional overall detection probability without the known radio collars (indicated by * superscript):

$$\hat{p}_{U,j,H(k)}^* = 1 - (1 - \hat{p}_{U,F,j,H(k)}) \cdot (1 - \hat{p}_{U,B,j,H(k)}) \quad (3)$$

These differ from those computed in Equation (2), which set the detection probability to 1 for any group with a radio collar. These probabilities represent our estimate of the unbiased (unconditional) probability that aerial observers alone would detect groups with the given covariate values. We computed conditional probabilities,

$\hat{p}_{C,j,H(k)}$, similarly. We then compared the model-averaged values:

$$\hat{p}_{C,j,H(\cdot)} = \sum_k \hat{w}_{H(k)} \hat{p}_{C,j,H(k)} \quad (4)$$

$$\hat{p}_{U,j,H(\cdot)}^* = \sum_k \hat{w}_{H(k)} \hat{p}_{U,j,H(k)}^* \quad (5)$$

where $\hat{w}_{H(k)}$ is the AIC_c weighting factor for model M_H , variant k (Burnham and Anderson 2002). The notation $H(\cdot)$ indicates that the values are averaged across the M_H models. For any given $x_{j,i}$, the heterogeneity bias was the difference between $\hat{p}_{C,j,H(\cdot)}$ and $\hat{p}_{U,j,H(\cdot)}^*$.

We also assessed model performance by comparing probabilities of detection estimated from the models M_H , M_S , and M_D . For a range of $x_{j,i}$ values, we compared the model-averaged overall unconditional detection probability, $\hat{p}_{U,j,H(\cdot)}^*$ from model M_H to the expectation of $\hat{p}_{C,j,D(\cdot)}$ from model M_D and to $\hat{p}_{U,j,S(\cdot)}$ from model M_S .

Population Abundance and Composition

M_H model estimates.—We estimated the number of animals in the survey area that had no radio-collars in the group and were not seen during the survey by weighting the counts of elk within observed groups by the inverse of the unconditional detection probabilities estimated for each observed group from sighting covariates. For each of the k M_H models, we computed specific correction factors, $\hat{\theta}_{j,H(k)}$, for each group, j , as $1/\hat{p}_{U,j,H(k)}^*$. The application of such Horvitz-Thompson correction factors for estimating detection biases in aerial surveys is well established (Steinhorst and Samuel 1989, Borchers et al. 1998). Group-specific correction factors derived from $\hat{p}_{U,j,H(k)}^*$ can be thought of as the per-observed-elk contribution to the overall estimate of abundance. We computed a model-averaged estimate of abundance:

$$\bar{N}_H = \sum_j \sum_k n_j \cdot \hat{\theta}_{j,H(k)} \cdot \hat{w}_{H(k)} \quad (6)$$

In Equation (6), n_j is the recorded number of elk observed in group j . We did not apply a correction factor to elk groups with missing covariate values. We added elk in such groups without correction to the model-based estimate.

For each survey, we estimated composition of the herds from the total estimated numbers of cows, calves, and bulls, by multiplying the observed numbers of cows, calves, and bulls in each group, j , by the group-specific correction factor:

$$\bar{N}_C = \sum_k \sum_j n_{j,c} \cdot \hat{\theta}_{j,H(k)} \cdot \hat{w}_{H(k)} \quad (7)$$

Here, \bar{N}_C is the estimated total number of elk in age-sex class c (c = cow, calf, or bull) and $n_{j,c}$ is the number of elk of age-sex class c seen in group j . We converted the estimated numbers of cows, calves, and bulls in the 9 surveys to ratios of bulls per 100 cows and calves per 100 cows, and then compared those ratios to values from uncorrected counts.

We estimated variance of abundance estimates from M_H using a data bootstrapping method (Wong 1996; see supplemental material, available online at www.onlinelibrary.

wiley.com) to account for variance due to 1) random detection of animal groups within surveyed areas, 2) variation in sightability model estimation (Steinhorst and Samuel 1989, Fieberg and Giudice 2008), and 3) model selection uncertainty. This variance estimation procedure has the added advantage of not depending on an assumption that group sizes are normally distributed. Because we searched survey areas entirely, we did not introduce variation due to sampling of survey subunits.

Sightability models and double-observer models.—For model M_S , we fit 6 Huggins (1989, 1991) models in Program MARK (White and Burnham 1999), fixing the probability for the telemetry observer equal to 1. Because of smaller sample size, our M_S models only considered the most highly supported covariates out of those in the M_H models considered (see Results); each model structure had effects of $\ln(n)$ and V , and at most 1 other covariate. We used the SightabilityModel package (Fieberg 2012) in the R statistical software (R Development Core Team 2010) to estimate sightability-based abundance and variance of abundance for each survey and each model, and then used sightability model weights to find model-averaged sightability-based estimates of abundance, \hat{N}_S , and variance of \hat{N}_S . Because we applied sightability models to survey data as is typically done after sightability model development, we did not include counts of elk groups with radio collars that were missed during survey in sightability-based abundance estimates. For a range of x_{ji} values, we compared group-specific detection probabilities from the model-averaged M_S models, $\hat{p}_{U,j,S(\cdot)}$, to $\hat{p}_{U,j,H(\cdot)}^*$ from the M_H model. We also compared model-averaged population estimates and coefficients of variation from the M_S model to the M_H model.

We parameterized 5 Huggins (1989, 1991) DO models with individual covariates (M_D models) fit to our DO observation data in Program MARK (White and Burnham 1999). We estimated model-averaged abundance, \hat{N}_D , for each survey. We estimated the variance of \hat{N}_D for each survey using the delta method (Seber 1982) applied to variance in the expected number of undetected groups in each survey and variance in the average expected group size of undetected groups. We compared estimated population estimates and coefficients of variation from the M_D model to the M_H model.

RESULTS

We recorded 510 usable DO observations and 97 DO-S trials. We also recorded but did not use 38 DO observations because of incomplete sighting covariates or lack of independence due to 1 set of observers alerting the other of their observations. We assumed the deleted points were a random subset of all data, as we found no pattern in omitted points that would have introduced bias.

Model Fitting and Selection

Of 22 M_H models considered, we found good support ($\Delta AIC_c < 4.0$) for 15 models (Table 3; see Supplemental Table S1, available online at www.onlinelibrary.wiley.com) and used these for multi-model inference (model averaging).

In addition to the variables included in all models a priori (i.e., $\ln(n)$, P , and C), we found strong support for including vegetation cover (V) in the model. No model without V was included among the 15 most strongly supported models. The influence of lighting (L), animal movement (M), herbaceous vegetation (H), and forest vegetation (F) were also included in some supported models, but we found less support for those covariates than for V based on evidence ratios (ER). ERs computed as the ratio AIC_c weights, \hat{w}_k , of the most supported model with an effect to the nested model without that effect, ranged from 88.1 for the effect of V (i.e., the evidence for inclusion of V was 88 times stronger than the evidence for excluding it), 1.4 for L , to <1.0 for F (0.37), H (0.5), and M (0.95). We found little support for separate sighting probabilities for front- and back-seat observers (Table 3; Supplemental Table S1; ER = 0.55) and only weak support for differences in the residual heterogeneity parameter between front- and back-seat observers (Table 3; Supplemental Table S1; ER = 1.1). Model averaging among the top 15 models reduced the already minor estimated differences in detection probabilities and biases between front- and back-seat observers.

Group size had a strong influence on detection probabilities, particularly for small groups (Fig. 2). Greater degrees of concealing vegetation cover were associated with lower detection probabilities (Fig. 3). Having an inexperienced pilot-observer reduced $\hat{p}_{U,F,j,H(\cdot)}$ by 0.15–0.17 if the elk group was on the pilot's side of the helicopter flight path. For elk groups along the flight path centerline, the back-seat observers' detection probability, $\hat{p}_{U,B,j,H(\cdot)}$, was nearly zero, which substantially reduced overall sighting probability, $\hat{p}_{U,j,H(\cdot)}^*$, for those groups.

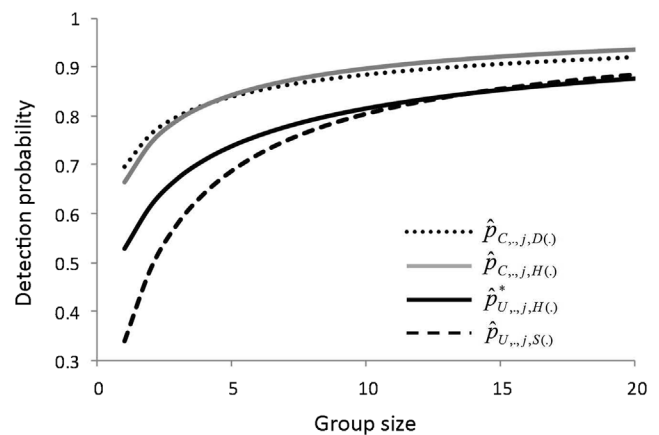


Figure 2. Expected model-averaged elk group detection probabilities for aerial elk surveys in Mount Rainier National Park, Washington, USA, 2008–2010. We show predicted detection probabilities as a function of group size (x -axis). We present the unconditional detection probability for elk group j from the M_H model (solid black line, $\hat{p}_{U,j,H(\cdot)}^*$), and the biased estimate of detection probability from the M_H model (gray line; $\hat{p}_{C,j,H(\cdot)}$). Estimates of the biased probabilities do not include the corrective effect of the heterogeneity bias parameter. We also show the model-averaged detection probabilities from model M_D (dotted line, $\hat{p}_{C,j,D(\cdot)}$) and model M_S (dashed line, $\hat{p}_{U,j,S(\cdot)}$). We show results for flights with an experienced pilot, under flat lighting conditions, for a stationary elk group in forested substrate of 26–50% concealing vegetation cover.

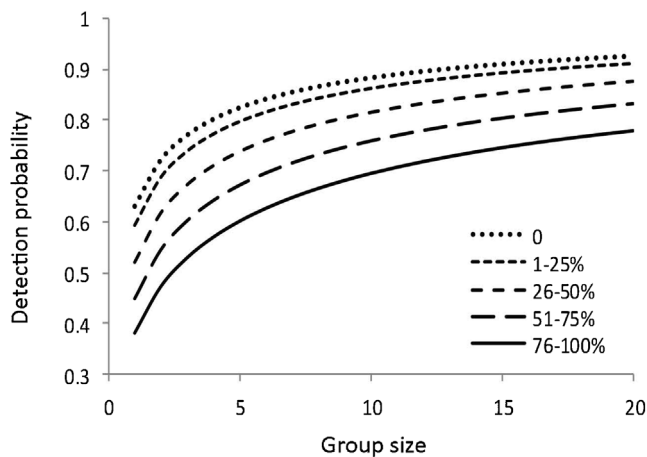


Figure 3. Expected model-averaged unconditional detection probabilities, $\hat{p}_{U,j,H(\cdot)}^*$, as a function of group size (x -axis) and 5 values of percentage concealing vegetation (curves), from a model M_H for elk aerial surveys in Mount Rainier National Park, Washington, USA, 2008–2010. We show calculations for flights with an experienced pilot, under flat lighting conditions, for a stationary elk group in forested cover. We used the midpoint of the recorded ranges (0%, 13%, 38%, 63%, and 88%) as the percent concealing vegetation values.

All 6 of the candidate M_S models considered had good support ($\Delta AIC_c < 4.0$) and were used in model averaging (Table 3; Supplemental Table S1). Of the 5 optional covariates considered, lighting, L , had the most support in the M_S models ($ER = 1.2$). Similarly, all 5 candidate M_D models received sufficient support to be used in multi-model inferences (Table 3; Supplemental Table S1). Of the 4 optional covariates considered in the M_D models, movement, M , had the most support ($ER = 0.9$).

Probability of Detection

Detection probabilities derived from the model-averaged M_H model, $\hat{p}_{U,j,H(\cdot)}^*$, were consistently lower than the biased probability that did not include the heterogeneity parameter, $\hat{p}_{C,j,H(\cdot)}$ (Fig. 2). The difference between $\hat{p}_{U,j,H(\cdot)}^*$ and $\hat{p}_{C,j,H(\cdot)}$, which represented the magnitude of heterogeneity

bias estimated from M_H models, averaged approximately 5–10% across groups.

Overall expected detection probabilities from model M_D , $\hat{p}_{C,j,D(\cdot)}$, were comparable to estimates from the M_H model before correcting for bias, $\hat{p}_{C,j,H(\cdot)}$. Expected detection probabilities from the sightability model, $\hat{p}_{U,j,S(\cdot)}$, were close to unbiased estimates, $\hat{p}_{U,j,H(\cdot)}^*$, from model M_H for group size >10 elk, but lower for smaller groups (Fig. 2).

Population Abundance and Composition

Model M_D led to estimates of abundance averaging 4.2% lower than M_H (2-tailed paired t -test, $df = 8$, $P = 0.005$) and nearly identical precision (2-tailed paired t -test, $df = 8$, $P = 0.47$; Table 4). Model M_S produced abundance estimates that were, on average, 4.1% greater than those from Model M_H (2-tailed paired t -test, $df = 8$, $P = 0.099$) but with much lower precision than estimates from Model M_H (2-tailed paired t -test, $df = 8$, $P = 0.024$).

Application of the model M_H increased estimated ratios of bulls per 100 cows, compared to ratios from uncorrected counts (Table 5; mean increase = 5.0, $SE = 1.2$; 2-tailed paired t -test, $df = 8$, $P = 0.003$). Estimated ratios of calves per 100 cows were nearly equal to ratios from uncorrected counts (Table 5; mean difference = 0.05, $SE = 0.40$; 2-tailed paired t -test, $df = 8$, $P = 0.89$).

DISCUSSION

In comparing uncorrected counts to abundance estimates from model M_H , the survey crews detected 80–93% of elk present in Mount Rainier National Park subalpine parklands. Elk sightability during helicopter surveys was determined primarily by group size, the amount of concealing vegetation cover, and, to a lesser extent, animal movement and light level; all are factors that have been identified previously in sightability models of detection bias in elk aerial surveys (e.g., Samuel et al. 1987, Anderson et al. 1998, Gilbert and Moeller 2008, McIntosh et al. 2009). Model M_H also allowed us to quantify the negative effect of inexperienced

Table 4. Uncorrected counts and estimated abundance values for elk surveys within North Rainier (north) and South Rainier (south) trend count areas, Mount Rainier National Park, Washington, USA, 2008–2010. Estimates of abundance (\hat{N}) and standard errors (SE) reflect application of 3 modeling approaches: hybrid double-observer sightability (model M_H), sightability (model M_S), and double-observer (model M_D). Coefficients of variation (CV) are the standard errors divided by the point estimate of abundance.

Survey area, survey date	Raw count	Model M_H			Model M_S			Model M_D			Comparisons			
		N_H	SE_H	CV_H (%)	N_S	SE_S	CV_S (%)	N_D	SE_D	CV_D (%)	Estimates		CV	
											$\hat{N}_S - \hat{N}_H$ (%)	$\hat{N}_D - \hat{N}_H$ (%)	$\frac{CV_S - CV_H}{CV_H}$ (%)	$\frac{CV_D - CV_H}{CV_H}$ (%)
North, 22–26 Aug 2008	214	261.8	18.0	6.9	279.4	34.2	12.2	246.1	20.9	8.5	6.7	–6.0	78.0	23.5
North, 3–6 Sep 2008	240	294.2	29.3	10.0	318.1	38.2	12.0	276.8	23.5	8.5	8.1	–5.9	20.6	–14.8
South, 10–11 Sep 2008	314	391.3	30.5	7.8	405.8	42.9	10.6	356.4	26.3	7.4	3.7	–8.9	35.6	–5.3
North, 17 Aug 2009	363	424.4	21.4	5.0	460.6	45	9.8	425	36.1	8.5	8.5	0.1	93.8	68.5
South, 18 Aug 2009	397	491.1	62.5	12.7	480.8	44.2	9.2	475.1	38.4	8.1	–2.1	–3.3	–27.8	–36.5
South, 2 Sep 2009	220	268.5	25.4	9.5	306.9	51	16.6	250.7	20.9	8.3	14.3	–6.6	75.7	–11.9
North, 17 Aug 2010	288	310.6	12.8	4.1	317.5	24.2	7.6	311.3	15.5	5.0	2.2	0.2	85.0	20.8
South, 18 Aug 2010	615	707.1	49.3	7.0	715.1	52.7	7.4	670.9	36.8	5.5	1.1	–5.1	5.7	–21.3
North, 25 Aug 2010	375	410.2	26.6	6.5	387.3	37.5	9.7	399	17.6	4.4	–5.6	–2.7	49.3	–32.0
Mean											4.1	–4.2	46.2	–1.0
SE											2.0	1.0	13.7	11.1

Table 5. Estimated corrected and uncorrected composition of elk within the North Rainier (north) or South Rainier (south) survey areas, Mount Rainier National Park, Washington, USA, 2008–2010. Estimates of composition reflect application of group-specific correction factors that result from the hybrid double-observer sightability model, whereas uncorrected values only reflect the tally of observed cows, calves, and bulls.

Survey area, survey date	Uncorrected calves: 100 cows	Estimated calves: 100 cows	Uncorrected bulls: 100 cows	Estimated bulls: 100 cows
North, 22–26 Aug 2008	49.6	47.1	32.8	34.4
North, 3–6 Sep 2008	44.4	45.6	42.1	45.2
South, 10–11 Sep 2008	33.7	34.1	34.8	36.1
North, 17 Aug 2009	37.9	37.1	21.5	31.5
South, 18 Aug 2009	34.7	34.0	30.1	32.6
South, 2 Sep 2009	32.4	32.5	22.3	27.9
North, 17 Aug 2010	33.3	33.6	42.0	53.2
South, 18 Aug 2010	31.7	31.7	29.4	33.4
North, 25 Aug 2010	48.7	50.2	40.5	46.6

pilot-observers on the front-seat observer pair's detection probability, and the negative effect of elk presence along the flightline center on detection probabilities of back-seat observers.

Model M_H allowed for the explicit estimation of heterogeneity biases. These are undercounting biases that result from co-dependencies in sighting probabilities by front- and back-seat observers, particularly the tendencies of front- and back-seat observers to both miss or to both see the same groups because of the characteristics of those groups. A simplistic DO model with no covariates would be biased because of such heterogeneity. In DO models, adding covariates may reduce heterogeneity in detection probabilities, but still requires the assumption that all groups with identical covariates have identical sighting probability. However, including all covariates that affect sighting probability is not possible. Nor can we know what factors not included in the candidate models might have been influential. For example, groups at a greater distance from observers are almost certainly less visible, but distance is difficult to measure accurately in steep terrain and we did not record it during our surveys. Model M_H adds an additional bias correction to capture the average effect of all of the residual factors that we did not explicitly measure and include in the model.

Model M_H used data from missed, radio-collared animals to quantify the otherwise unmodeled component of heterogeneity bias. The estimation of heterogeneity bias in model M_H reduced detection probabilities in comparison to model M_D . Model M_D overestimated detection probabilities and underestimated abundance relative to model M_H by 4.2% in the relatively open subalpine environments of Washington's Cascades Range. This relatively small residual bias indicates that our selected covariates successfully modeled most of the heterogeneity in sighting probabilities among groups. The effect of bias due to unmodeled heterogeneity might be greater in aerial surveys where vegetative cover is more dense or surveyed animals more difficult to see because the magnitude of this bias in DO analyses generally increases as the unconditional detection probability decreases (Barker 2008).

We predicted that abundance estimated from model M_H would be more precise than estimates derived from standard sightability modeling, M_S , because model M_H draws from a

larger data set. In our surveys, approximately 17% of the observations used to develop the model contained radio-collared animals. Consistent with the prediction, coefficients of variation of abundance from model M_H were significantly and substantially less than comparable estimates from model M_S . The improved precision under model M_H could be useful if management applications require narrow confidence intervals or more immediate detection of temporal changes in abundance. If M_H models are developed for other species or other geographic areas, the improved precision expected from the hybrid approach over the sightability approach will depend on the sample size of observations included in model fitting. The precision of the heterogeneity bias parameter estimate is affected by the sample size of DO-S trials, whereas the precision of all other coefficients can be improved by increasing the sample size of all observations, including DO observations, which continue to accumulate with each new survey.

Applying model M_H to estimate composition led to greater ratios of bulls per 100 cows than if we had computed the ratio from uncorrected counts. This result supports previous conclusions that biases in bull/cow ratios may result from the tendency for bulls to occur alone or in smaller, less easily detected groups (McCorquodale 2001). Because we had no radio-marked bull elk, we cannot discern if sightability factors differed between bulls and cows independent from the effects of group size and vegetation, although that may be an interesting area for future research. In contrast to the influence of sightability biases on sex ratios, we found no discrepancies in calf-cow ratios derived from raw counts versus those derived from any of the sightability modeling approaches. We concluded that estimates of cow/calf composition were relatively unaffected by detection biases because cows and calves generally are found in the same groups at the time of the surveys and hence are detected with similar probabilities.

One salient benefit of recording the pattern of independent observations by crew members is that these DO observations can be collected during ongoing surveys. An ever-growing data set can improve the accuracy of models fit to those data and allows for evaluation of temporal changes in conditional detection probabilities. Ideally, for M_H models to fully address all sources of heterogeneity bias, ongoing or periodic assessment of heterogeneity biases derived from uncondi-

tional sighting probabilities of radio-collared groups is desirable. However, given the expense of maintaining functional radio-collars in populations, it may be practical to apply M_H models to future surveys following a model development period, after radio-collars have expired, as is commonly done with M_S modeling approaches. Future applications of the model without the use of radio-collared animals require the implicit assumption that the heterogeneity bias correction is constant. This is the same limitation inherent in M_S after model development is completed and established models are applied in the future. Unlike M_S models, in which all factors influencing sightability must be assumed to remain constant, all but 1 factor in M_H models may be continuously updated and refined with additional DO data from each survey. Only the correction due to the implicit heterogeneity bias is not updated unless unconditional detection biases are re-estimated using radio-collars. Importantly, the contribution of this single factor to the correction is only a fraction of the total so any drift in heterogeneity bias over time will result in proportionally less error in future surveys relative to the error in M_S models. Nevertheless, we recommend using marked animals at regular intervals to update the calibration of the implicit heterogeneity bias parameter.

Previous studies have highlighted the potential effects of observer variability on detection biases (Samuel et al. 1987, Quang and Becker 1997, Walter and Hone 2003). Because of the number of observers participating in this multi-tribal, multi-agency program, we were not able to model individual observer acuity, as has been done in other studies (Nichols et al. 2000, Schoenecker et al. 2006, Lubow and Ransom 2007, Shirley et al. 2012). Instead, we chose to allow variance due to variable observer acuity to affect the uncertainty in coefficient estimates for model covariates. Despite this, the large sample size of observations led to fairly precise estimates of abundance for the model M_H , with coefficients of variation between 4.1% and 12.7% over 9 surveys. Moreover, assuming the range of acuities in the future is comparable to that which led to the original data set, model M_H derived from pooled observations from many experienced observers may be robust to changing survey crews in the future.

When a factor affecting detection probability is not included in a model, the variance of beta parameters in the model can be expected to increase, and the precision of the resulting abundance estimates to decrease. In cases where the percentage of groups seen from year to year is lower and more variable than in our study (e.g., Cogan and Diefenbach 1998, 43–85%; Anderson et al. 1998, 67–80%), the DO-S approach could be used to estimate a year effect to account for large annual changes in detection probability. In those cases, maintaining more radio-collared animals may be important, so one could test models with annual variation in heterogeneity bias.

The hybrid method recommended here is suitable for most aerial surveys of large terrestrial wildlife species. The field methods are only slightly more complex and generally no more costly than for either sightability or DO methods alone.

We do recognize, however, that the analytical methods are complex, particularly the bootstrap modeling required to estimate confidence intervals. To date, software to implement this has been developed specifically for each study and requires an experienced statistical analyst. Development of generalized software to facilitate analyses of this type is an important area for future work.

MANAGEMENT IMPLICATIONS

An unknown and variable number of animals go undetected in aerial surveys because of detection bias. The hybrid DO-S (M_H) model provides a flexible tool to adjust raw counts of animals observed during aerial surveys to obtain more accurate (both less biased and more precise) estimates of population abundance and composition. Unlike raw counts, the model also provides estimates of precision. Unlike standard DO (M_D) models, our approach quantifies the level of bias that comes from observer non-independence using data from radio-telemetered animal groups. Model M_H is suitable for areas where some animals are radio-marked, at least during some initial surveys, but a larger number are not. Model M_H tends to be more precise than standard sightability (M_S) modeling and less negatively biased than Model M_D . Model M_H has the added advantage over model M_S of accumulating data and refining the model following each monitoring survey, continuously increasing precision and allowing the model to track changes in sighting probability over time.

ACKNOWLEDGMENTS

This study was supported by the United States Geological Survey (USGS) National Park Monitoring Project, the NPS North Coast and Cascades Network, USGS Forest and Rangeland Ecosystem Science Center, Mount Rainier National Park, Olympic National Park, Washington's National Park Fund, the Muckleshoot Indian Tribe (MIT), the Puyallup Tribe of Indians (PTOI), and Washington Department of Fish and Wildlife (WDFW). We thank R. Lechleitner, S. Lofgren, and A. Robb for aviation management; pilots C. Cousins, J. Goodrich, J. Hagerman, R. Olmstead, and T. Walker; biologists and technicians A. Hise, E. Meyers (NPS), P. Dillon (PTOI), E. Anderson, T. Benson, M. Hilden, M. Middleton (MIT), A. Duff, E. Holman, S. Jonker, A. Prince, and T. Schmidt (WDFW) for help in surveys, P. Arnold for GIS assistance (PTOI); and J. Agee, P. Geissler, M. Huff, and P. Latham for review of early manuscripts. Any use of trade, product, or firm names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

LITERATURE CITED

- Anderson, C. R., and F. G. Lindzey. 1996. Moose sightability model developed from helicopter surveys. *Wildlife Society Bulletin* 24:247–259.
- Anderson, C. R., D. S. Moody, B. L. Smith, F. G. Lindzey, and R. P. Lanka. 1998. Development and evaluation of sightability models for summer elk surveys. *Journal of Wildlife Management* 62:1055–1066.

- Barker, R. 2008. Theory and application of mark-recapture and related techniques to aerial surveys of wildlife. *Wildlife Research* 35:268–274.
- Borchers, D. L., S. T. Buckland, P. W. Goedhart, E. D. Clarke, and S. L. Hedley. 1998. Horvitz-Thompson estimators for double-platform line transect surveys. *Biometrics* 54:1221–1237.
- Buckland, S. T., D. R. Anderson, K. P. Burnham, J. L. Laake, D. L. Borchers, and L. Thomas. 2004. *Advanced distance sampling*. Oxford University Press, Oxford, England.
- Buckland, S. T., J. L. Laake, and D. L. Borchers. 2010. Double-observer line transect methods: levels of independence. *Biometrics* 66:169–177.
- Burnham, K. P., and D. R. Anderson. 1984. The need for distance data in transect counts. *Journal of Wildlife Management* 48:1248–1254.
- Burnham, K. P., and D. R. Anderson. 2002. *Model selection and multimodel inference: a practical information-theoretic approach*. Second edition. Springer-Verlag, New York, New York, USA.
- Caughley, G., and D. Grice. 1982. A correction factor for counting emus from the air and its application to counts in Western Australia. *Australian Wildlife Research* 9:253–259.
- Cogan, R. D., and D. R. Diefenbach. 1998. Effect of undercounting and model selection on a sightability-adjustment estimator for elk. *Journal of Wildlife Management* 62:269–279.
- Crête, M., D. Vandal, L.-P. Rivest, and F. Potvin. 1991. Double counts in aerial surveys to estimate polar bear numbers during the ice-free period. *Arctic* 44:275–278.
- Fieberg, J. 2012. Estimating population abundance using sightability models: R SightabilityModel package. *Journal of Statistical Software* 51(9):1–20.
- Fieberg, J., and J. Giudice. 2008. Variance of stratified survey estimators with probability of detection adjustments. *Journal of Wildlife Management* 72:837–844.
- Franklin, J. F., W. H. Moir, M. A. Hemstrom, S. E. Greene, and B. G. Smith. 1988. *The forest communities of Mount Rainier National Park*. National Park Service Scientific Monograph, Series No. 19.
- Gilbert, B. A., and B. J. Moeller. 2008. Modeling elk sightability bias of aerial surveys during winter in the central Cascades. *Northwest Science* 82:222–228.
- Giudice, J. H., J. R. Fieberg, and M. S. Lenarz. 2012. Spending degrees of freedom in a poor economy: a case study of building a sightability model for moose in northeastern Minnesota. *Journal of Wildlife Management* 76:75–87.
- Graham, A., and R. Bell. 1989. Investigating observer bias in aerial survey by simultaneous double-counts. *Journal of Wildlife Management* 53:1009–1016.
- Huggins, R. M. 1989. On the statistical analysis of capture experiments. *Biometrika* 76:133–140.
- Huggins, R. M. 1991. Some practical aspects of a conditional likelihood approach to capture experiments. *Biometrics* 47:725–732.
- Krueger, B., L. C. Bender, W. R. Gould, and P. C. Morrow. 2007. A fixed-wing sightability model for oryx in desert habitats. *South African Journal of Wildlife Research* 2:133–142.
- Laake, J. L., B. A. Collier, M. L. Morrison, and R. N. Wilkins. 2011. Point-based mark-recapture distance sampling. *Journal of Agricultural, Biological, and Environmental Statistics* 16:389–408.
- Lubow, B. C., and J. I. Ransom. 2007. Aerial population estimates of wild horses (*Equus caballus*) in the Adobe Town and Salt Wells Creek Herd Management Areas using an integrated simultaneous double-count and sightability bias correction technique. U.S. Geological Survey Open-File Report 2007–1274. U.S. Geological Survey, Reston, Virginia, USA.
- Manning, J. A., and E. O. Garton. 2011. A sightability model for correcting visibility and availability biases in standardized surveys of breeding burrowing owls in southwest agroecosystem environments. *Journal of Wildlife Management* 76:65–74.
- Marsh, H., and D. F. Sinclair. 1989. Correcting for visibility bias in strip transect aerial surveys of aquatic fauna. *Journal of Wildlife Management* 53:1017–1024.
- McCorquodale, S. M. 2001. Sex-specific bias in helicopter surveys of elk: sightability and dispersion effect. *Journal of Wildlife Management* 65:216–225.
- McIntosh, T. E., R. C. Rosatte, J. Hamr, and D. L. Murray. 2009. Development of a sightability model for low-density elk populations in Ontario, Canada. *Journal of Wildlife Management* 73:580–585.
- Moeller, B. J. 2010. *Migratory and non-migratory movements and habitat use by female elk in the Cascade Range of Washington*. Thesis, Evergreen State College, Olympia, Washington, USA.
- Nichols, J. D., J. E. Hines, J. R. Sauer, F. W. Fallon, J. E. Fallon, and P. J. Heglund. 2000. A double-observer approach for estimating detection probability and abundance from point counts. *Auk* 117:393–408.
- Pacific Meridian Resources. 1996. *Vegetation and landform database development study: final report*. Pacific Meridian Resources, Portland, Oregon, USA.
- Pollock, K. H., and W. L. Kendall. 1987. Visibility bias in aerial surveys: a review of estimation procedures. *Journal of Wildlife Management* 51:502–510.
- Pollock, K. H., H. D. Marsh, I. R. Lawler, and M. W. Alldredge. 2006. Estimating animal abundance in heterogeneous environments: an application to aerial surveys for dugongs. *Journal of Wildlife Management* 70:255–262.
- Potvin, F., L. Breton, and L.-P. Rivest. 2004. Aerial surveys for white-tailed deer with the double-count technique in Québec: two 5-year plans completed. *Wildlife Society Bulletin* 32:1099–1107.
- Quang, P. X., and E. F. Becker. 1997. Combining line transect and double count sampling techniques for aerial surveys. *Journal of Agricultural, Biological, and Environmental Statistics* 2:230–242.
- R Development Core Team. 2010. R: a language and environment for statistical computing, Version 2.12.0. R Foundation for Statistical Computing, Vienna, Austria: ISBN 3-900051-07-0. < <http://www.R-project.org/> >. Accessed 18 Apr 2012.
- Ransom, J. I. 2012. Detection probability in aerial surveys of feral horses. *Journal of Wildlife Management* 76:299–307.
- Rice, C. G., K. J. Jenkins, and W.-Y. Chang. 2008. A sightability model for mountain goats. *Journal of Wildlife Management* 73:468–478.
- Samuel, M. D., E. O. Garton, M. W. Schlegel, and R. G. Carson. 1987. Visibility bias during aerial surveys of elk in Northcentral Idaho. *Journal of Wildlife Management* 51:622–630.
- Samuel, M. D., and K. H. Pollock. 1981. Correction of visibility bias in aerial surveys where animals occur in groups. *Journal of Wildlife Management* 45:993–997.
- Schoenecker, K., B. Lubow, L. Ziegenfuss, and J. Mao. 2006. 2005 Annual progress report: elk and bison grazing ecology in the Great Sand Dunes Complex of Lands. U.S. Geological Survey Open-File Report 2006–1267. U.S. Geological Survey, Reston, Virginia, USA.
- Seber, G. A. F. 1982. *The estimation of animal abundance and related parameters*. Second edition. Macmillan, New York, New York, USA.
- Shirley, M. H., R. M. Dorazio, E. Abassery, A. A. Elhady, M. S. Mekki, and H. H. Asran. 2012. A sampling design and model for estimating abundance of Nile crocodiles while accounting for heterogeneity of detectability of multiple observers. *Journal of Wildlife Management* 76:966–975.
- Skalski, J. R., J. J. Millsbaugh, and R. D. Spencer. 2005. Population estimation and biases in paintball, mark-resight surveys of elk. *Journal of Wildlife Management* 69:1043–1052.
- Southwell, C., D. Borchers, C. G. M. Paxton, L. Burt, and W. de la Mare. 2007. Estimation of detection probability in aerial surveys of Antarctic pack-ice seals. *Journal of Agricultural, Biological, and Environmental Statistics* 12:41–54.
- Steinhorst, R. K., and M. D. Samuel. 1989. Sightability adjustment methods for aerial surveys of wildlife populations. *Biometrics* 45:415–425.
- Udevitz, M. S., B. S. Shults, L. G. Adams, and C. Kleckner. 2006. Evaluation of aerial survey methods for Dall's sheep. *Wildlife Society Bulletin* 34:732–740.
- Unsworth, J. W., F. A. Leban, D. J. Leptich, E. O. Garton, and P. Zager. 1999. *Aerial survey: user's manual with practical tips for designing and conducting aerial big game surveys*. Idaho. Department of Fish and Game, Boise, Idaho, USA.
- Walter, M. J., and J. Hone. 2003. A comparison of 3 aerial survey techniques to estimate wild horse abundance in the Australian Alps. *Wildlife Society Bulletin* 31:1138–1149.

- White, G. C. 1996. Noremark: population estimation from mark-resighting surveys. *Wildlife Society Bulletin* 24:50–52.
- White, G. C., and K. P. Burnham. 1999. Program MARK: survival estimations from populations of marked animals. *Bird Study* 46 (Supplement):120–138.
- Williams, B. K., J. D. Nichols, and M. J. Conroy. 2002. Analysis and management of animal populations: modeling, estimation, and decision making. Academic Press, San Diego, California, USA.
- Wong, C. 1996. Population size estimation using the modified Horvitz-Thompson estimator with estimated sighting probabilities. Dissertation, Colorado State University, Fort Collins, USA.

Associate Editor: David Forsyth.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of this article at the publisher's web-site.

Table S1. Matrix of ΔAIC_c values, AIC_c model weights (w_k), and β coefficient estimates for each of the hybrid double-observer model structures (Model M_H), sightability model structures (model M_S), and double-observer model structures (model M_D) developed using elk aerial survey data from Mount Rainier NP, Washington, USA.